(Part VI) Meta-genesis. Towards a biology without matter, based on pure logic

# Multi-Scale Numerical Invariants and Fractal Properties of the Genetic Code: A Combinatorial and Atomic Analysis

Victoria Kayser-Cuny

**Keywords:** quantum biology, biomathematics, bioinformatics, information theory, genetic code, combinatorics, invariants, codons, DNA, atoms, numerical scales, mathematical constants, astrobiology, exobiology.

**Abstract**

The genetic code is based on an alphabet of four bases (A, U, C, G), organized into triplets called codons. By systematically analyzing the combinations at the levels of nucleotides, theoretical duplets (1), and triplets, and then applying a simple numerical transformation — $n^3/4$ —, I highlighted the recurrent emergence of specific constants (1, ≈96, 128). These values also appear at the atomic level for the constituent elements of DNA (C, H, O, N), suggesting that they do not result from chance, but from fundamental combinatorial constraints linked to the quaternary alphabet and the chemical composition of the bases.

This reveals the existence of multi-scale numerical invariants comparable to fractal patterns, without implying any direct biological function. The scientific interest of this discovery does not lie in a new biological application, but in the perspective it opens: considering the genetic code not only as a sequence of biochemical information, but also as a mathematical object.

The poetic analogy is illuminating: a poem is generally studied for its meaning and emotions, but it also obeys formal rules (syllables, rhymes, meters). My approach consists of unveiling these invisible rules in the genome. As in a Sierpinski triangle, where a pattern repeats infinitely, the same constants (1, ≈96, 128) manifest at different scales, from the atom to the codon.

This organization reveals that the genetic code is not random: it is structured by universal arithmetic constraints. These invariants may represent the fundamental building blocks that guided its emergence and evolution. The multi-scale numerical invariants identified (1, ≈96, 128) suggest that the structure of the genetic code is not a purely contingent product of terrestrial evolution, but the result of combinatorial constraints imposed by the chemistry of the four main constitutive elements of DNA (C, H, O, N). These elements possess fixed numbers of protons, neutrons, and electrons, which, when combined in nucleotides and codons, generate universal arithmetic regularities. Thus, DNA can be interpreted not only as a biochemical support for information, but also as the logical consequence of the atomic combinatorics of CHON.

This perspective opens a theoretical framework in which the genetic code appears as a quasi-necessary structure, whose organization derives directly from the fundamental properties of organic chemistry. It suggests that the emergence of DNA, or of a functionally analogous molecule, could be inevitable whenever life based on CHON appears, regardless of the geological or evolutionary context.

In this sense, my work belongs to the field of theoretical biomathematics and bioinformatics: it does not aim to explain a new function, but to shed light on the hidden elegance of the structure of life.

This study extends a previous work (*The Mirror-Twin Paradox: A New Approach to DNA*), where I proposed an inversion transformation (A↔G, C↔T) generating a “mirror genome” — a symmetry — tested on genetic databases (2). In both cases, the objective remains the same: to apply a systematic mathematical transformation to DNA in order to reveal structures invisible to a classical biological reading. Together, these approaches reinforce the idea that the genome possesses a universal internal organization, which can be revealed through operations of symmetry or numerical transformation.

**Introduction**

The genetic code is described as universal and redundant: 64 codons, composed of triplets of four bases, encode 20 amino acids and three stop signals. Although variants exist (33 codes identified by the NCBI), the basic structure remains that of a quaternary alphabet combined in triplets.

The present study explores this code from an arithmetic and combinatorial perspective. By applying a simple numerical transformation — $n^3/4$ —, we examined the distribution of the values obtained at several scales:

- constituent atoms of DNA (C, H, O, N),
- nucleotidic bases,
- theoretical doublets that do not exist in nature (16 combinations),
- triplets (64 codons).

We show that recurrent constants (1, ≈96, 128) emerge at each of these scales, and discuss their origin.

**Methodology**

Unlike the classical approach, which considers DNA as a mere carrier of chemical information, this study treats it as a self-organized logical system. Each codon (triplet of nitrogenous bases) is converted into a *theoretical codonic unit* (TCU) derived from the product of the number of protons, neutrons, and electrons present in the constituent atoms (C, H, N, O, excluding phosphorus). This product yields a kind of “logical mass” or state

volume, which is then normalized and translated into binary form. This approach does not alter the actual chemistry: it simply reveals, through a mathematical magnification — like a lens — patterns already inscribed in the atomic structure.

$$v_i = p_i \cdot n_i \cdot e_i$$

**v** = logical value of a codon raised to the cube
**p** = number of protons
**n** = number of neutrons
**e** = number of electrons

The value, once cubed, is then divided by 4 in order to reflect the quaternary nature of the system (four main atoms, four bases, codons organized into groups of four):

$$f(n) = \frac{n^3}{4}$$

This methodology brings out constant numerical patterns that appear to correspond to logical transitions. The underlying hypothesis is that the architecture of DNA encodes ON/OFF-type states within a logical sequence, organizing life as a natural finite automaton.

## PROPOSAL FOR A THEORETICAL CODONIC UNIT (TCU)

| | Series of numbers – subatomic particles raised to the cube | Combinaison | Family | Number of particles cubed and divided by 4 | Constant numerical patterns |
|---|---|---|---|---|---|
| | 1 | Electron | Particle | 0.25 | Does not apply. |
| +26 | 27 (3 quarks) | Proton | Particle | 6.75 | |
| +0 | 27 (3 quarks) | Neutron | Particle | 6.75 | |
| | 1 | Hydrogène | Atom | 0.25 ou 1 ? | **1** |
| +215 | 216 | Carbone | Atom | 54 | +54 |
| +127 | 343 | Azote | Atom | 85.75 | +31.75 |
| +169 | 512 | Oxygène | Atom | 128 | +42.25 (note : 54 + 42.25 = |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | **96.25**, et 54 + 31.75 + 42.25 = **128**) |
| +1898 | 2410 | C | Nucleic base | 602.5 | |
| +386 | 2796 | T | Nucleic base | 699 | **+96.5** |
| +4 | 2800 | A | Nucleic base | 700 | **+1** |
| +512 | 3312 | G | Nucleic base | 828 | **+128** |
| +1508 | 4820 | CC | Theoretical ancient code with doublets | 1205 | |
| +386 | 5206 | TC-CT | Theoretical ancient code with doublets | 1301.5 | **+96.5** |
| +4 | 5210 | AC-CA | Theoretical ancient code with doublets | 1302.5 | **+1** |
| +382 | 5592 | TT | Theoretical ancient code with doublets | 1398 | +95.5 |
| +4 | 5596 | AT-TA | Theoretical ancient code with doublets | 1399 | +1 |
| +4 | 5600 | AA | Theoretical ancient code with doublets | 1400 | +1 |
| +122 | 5722 | CG-GC | Theoretical ancient code with doublets | 1430.5 | +30.5 (note : 95.5 + 1 + 1 + 30.5 = **128**) |
| +386 | 6108 | TG-GT | Theoretical ancient code with doublets | 1527 | **+96.5** |
| +4 | 6112 | AG-GA | Theoretical ancient code with doublets | 1528 | **+1** |
| +512 | 6624 | GG | Theoretical ancient code with doublets | 1656 | **+128** |
| +606 | 7230 | CCC | Codon | 1807 | |

| | | | | | |
|---|---|---|---|---|---|
| +386 | 7616 | CCT-TCC-CTC | Codon | 1904 | **+97** |
| +4 | 7620 | CCA-ACC-CAC | Codon | 1905 | **+1** |
| +382 | 8002 | TCT-CTT-TTC | Codon | 2000 | +95 |
| +4 | 8006 | TCA-ACT-CTA-ATC-CAT-TAC | Codon | 2001 | +1 |
| +4 | 8010 | ACA-CAA-AAC | Codon | 2002 | +1 |
| +122 | 8132 | CCG-GCC-CGC | Codon | 2033 | +31 (note : 95 + 1 + 1 + 31 = **128**) |
| +256 | 8388 | TTT | Codon | 2097 | +64 |
| +4 | 8392 | TTA-ATT-TAT | Codon | 2098 | +1 |
| +4 | 8396 | ATA-TAA-AAT | Codon | 2099 | +1 |
| +4 | 8400 | AAA | Codon | 2100 | +1 |
| +518 | 8518 | TCG-GCT-CTG-GTC-CGT-TGC | Codon | 2129 | +28 (note : 64 + 1 + 1 +1 +28 = **95**) |
| +4 | 8522 | ACG-GCA-CAG-GAC-CGA-AGC | Codon | 2130 | **+1** |
| +382 | 8904 | TTG-GTT-TGT | Codon | 2226 | +96 |
| +4 | 8908 | ATG-GTA-TAG-GAT-TGA-AGT | Codon | 2227 | +1 |
| +4 | 8912 | AAG-GAA-AGA | Codon | 2228 | +1 |
| +122 | 9034 | GCG-CGG-GGC | Codon | 2258 | +30 (note : 96 + 1 +1 + 30 = **128**) |

| | | | | | |
|---|---|---|---|---|---|
| +386 | 9420 | GTG-TGG-GGT | Codon | 2355 | **+97** |
| +4 | 9424 | GAG-AGG-GGA | Codon | 2356 | **+1** |
| +512 | 9936 | GGG | Codon | 2484 | **+128** |

**Particles (3+1):** $p^+$, $n^0$, $e^-$, + photon/ion (not in the table) but can correct the exception.

**Atoms (4):** C, H, O, N (phosphorus is considered a “peripheral addition”).

**Nucleic bases (4):** A, T, C, G.

**Doublets (16):** all possible combinations of two bases.

**Codons (64):** triplets, the core of the genetic code.

**Amino acids (20 + 3):** derived from the 64 codons, with a small “symmetry break” (redundancy of the code).

**Key idea**

All of DNA unfolds in multiples of 4 → a fractal and almost “digital” logic, from particles to the genetic language. The only apparent exception (3 particles) is resolved by adding a fourth invisible principle (photon or ion), which restores symmetry.

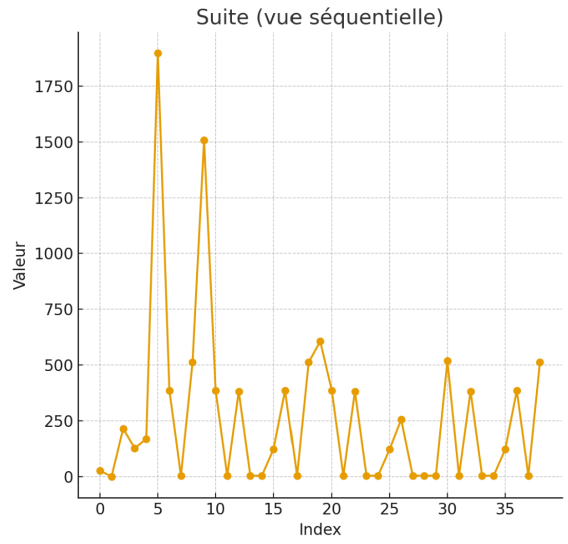
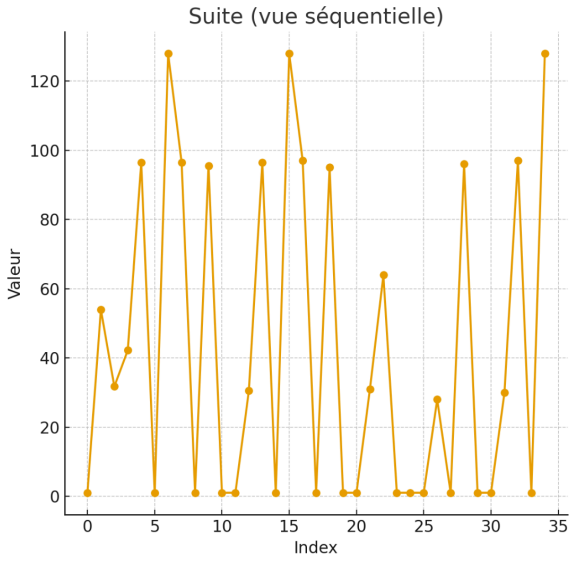
It is also possible to include the following codons (uracil instead of thymine):

| | | | |
|---|---|---|---|
| UGA | STOP | 8872 | 2218 |
| UAG | STOP | 8872 | 2218 |
| UAA | STOP | 8360 | 2090 |
| AUG | START | 8872 | 2218 |

**UGA, UAG, and AUG** have exactly the same logical value. This suggests that they are functional mirrors within the system, even though biologically they trigger opposite functions (START versus STOP). **UAA** is “softer,” weaker. There exists a logical polarity relationship in which starting (START) or stopping (STOP) amounts to inverting the same wave, like a transition with two faces of the same threshold.

These codons mark and structure the system's entry and exit points. One could describe them as phase vectors, logical gates within a multi-state system. It is a dynamic logical sequence structured by thresholds, loops, programmed interruptions, and formally defined transition states.

Are the missing numbers the intergenic spaces? If we consider that the system starts and stops at the same frequency, then it could possibly be seen as a closed oscillatory signal system!

<table>
<tr>
<td> 
</td>
<td> 
</td>
</tr>
<tr>
<td>1st column of the table: +26 +0 +215 +127 +169 +1898 +386 +4 +512 +1508 +386 +4 +382 +4 +4 +122 +386 +4 +512 +606 +386 +4 +382 +4 +4 +122 +256 +4 +4 +4 +518 +4 +382 +4 +4 +122 +386 +4 +512</td>
<td>6th column of the table: 1 +54 +31.75 +42.25 (note : 54 + 42.25 = 96.25, et 54 + 31.75 + 42.25 = 128) +96.5 +1 +128 +96.5 +1 +95.5 +1 +1 +30.5 (note : 95.5 + 1 + 1 + 30.5 = 128) +96.5 +1 +128 +97 +1 +95 +1 +1 +31 (note : 95 + 1 + 1 + 31 = 128) +64 +1 +1 +1 +28 (note : 64 + 1 + 1 +1 +28 = 95) +1 +96 +1 +1 +30 (note : 96 + 1 +1 + 30 = 128) +97 +1 +128</td>
</tr>
</table>

Compared numerical sequences from the codonic unit table proposed above.

<u>1st column of the table:</u>

In this numerical sequence, a central pattern emerges: the invariant block of 902, which can be written in two equivalent ways. In compact form, it appears as [386, 4, 512], while its expanded form is [386, 4, 382, 4, 4, 122], both leading to the same sum (386 + 4 + 512 = 902; 386 + 4 + 382 + 4 + 4 + 122 = 902). This duality is based on the identity 512 = 382 + 4 + 4 + 122, which reveals a logic of expansion/condensation within the system. Surrounding this recurrent block are insertions such as 1508, 606, 256, or 518, which, far from being arbitrary, are also anchored in an arithmetic structured by powers of two and multiples of 128. Thus, 386 and 382 appear as symmetric values around 384 (3 × 128), while 122 (128 − 6) and 518 (4 × 128 + 6) manifest another symmetry around 128. Altogether, this

suggests a numerical grammar based on powers of two (4, 256, 512) and ±2 or ±6 modulations around multiples of 128, organizing the sequence into an alternation of invariant blocks and modulatory variations.

6th column of the table:

Here, the organizing constant is no longer 902, but 128, which functions as a threshold or frame of reference. The sequence unfolds into a series of additive blocks whose sum systematically converges to 128, through slightly different combinations. For example, we observe equivalences such as 54 + 31.75 + 42.25 = 128, or 95.5 + 1 + 1 + 30.5 = 128, 95 + 1 + 1 + 31 = 128, 96 + 1 + 1 + 30 = 128, or even 97 + 30 + 1 = 128. In these blocks, the number 1 plays a dual role: it serves as a filler unit to adjust the sum precisely, but also as a separator between blocks. The recurrent appearance of values close to 96 (95, 95.5, 96, 96.5, 97) associated with complements (30, 30.5, 31, 31.75, 42.25) reflects a logic of symmetry around 96, while particular cases (64 + 1 + 1 + 1 + 28 = 95) indicate more elaborate variations. At several points, the total 128 also appears directly, as if to mark a “full” block. Altogether, this suggests a grammar based on the invariant sum 128, which unfolds either through adjusted combinations of 1 or 2 units, or through compact blocks where 128 is taken as is.

In this sequence, the logic is therefore no longer the equivalence between two representations of the same block (as for 902), but rather a multiplicity of paths to reach 128, either by decomposition or as a compact block.

Comparison between the two columns:

The comparison between the two sequences reveals that they belong to the same grammatical family, but at two different scales. In the first sequence, the invariant block is 902, which can be condensed into [386, 4, 512] or expanded into [386, 4, 382, 4, 4, 122], according to the internal identity 512 = 382 + 4 + 4 + 122. In the second sequence, the invariant is no longer 902 but 128, which unfolds into multiple additive combinations (95.5 + 1 + 1 + 30.5, 97 + 30 + 1, 96 + 1 + 1 + 30, etc.), or appears directly as a compact block. In both cases, the logic rests on a constant sum (902 or 128), around which the numbers aggregate according to a stable syntax: alternation between compact and expanded blocks, the role of units (1 or 4) as separators and adjusters, and numerical symmetries around 128 and its multiples (384, 512). Altogether, this outlines a modular architecture in which each sequence functions as a variation on the same principle: the stability of an invariant sum expressed through a diversity of arithmetical representations.

**A Boolean framework**

Boolean algebra is the foundation of electronic circuits composed of bistable devices—relays, switches, diodes, vacuum tubes, transistors—each of which can be

in an open (0) or closed (1) state. In this context, “1” corresponds to a closed circuit (minimal resistance), while “0” corresponds to an open circuit (maximal resistance). This logic allows complex switching networks to be simplified and efficiently analyzed.

In the living world, a striking parallel emerges: each cell nucleus contains a complete set of chromosomes—about 22,000 genes in humans—and operates through electrochemical phenomena. Each gene can be considered as a binary switch:

- **1 = ON** (gene expressed)
- **0 = OFF** (gene silent)

Thus, the genome behaves like a vast series circuit of molecular switches. A gene may carry a mutation—for example, the one responsible for hereditary hemochromatosis—while still remaining OFF, its effect remaining inactive as long as regulation does not turn it on.

Remarkably, the 64 codons of the genetic code lend themselves perfectly to this Boolean interpretation. Just as electronic engineers apply the first law of Boolean algebra to optimize an on/off circuit, one can view the codon table as a logical schema: each triplet acts as a binary input pattern, and the corresponding amino acid is the deterministic “output” of this logical operation.

> *“Boolean algebra, applied to electronics, deals with two-state devices such as relays, switches, diodes, vacuum tubes, and transistors. It allows rapid and efficient simplification of switching circuits. The two fundamental states are short-circuit – open, on – off, etc. In many applications, 1 = ON (short-circuit or minimal resistance) and 0 = OFF (open or maximal resistance).”*
> — Angelo C. Gillie (3)

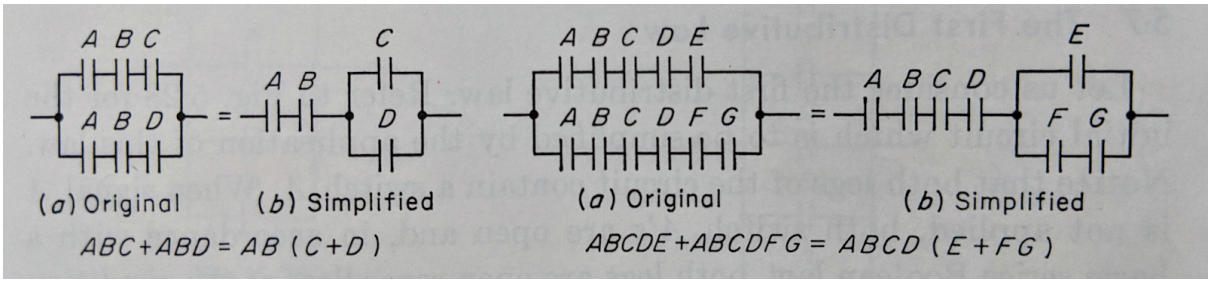



Illustration of the First Distributive Law
Binary Arithmetic and Boolean Algebra, p. 82 (Angelo C. Gillie, 1965).

| DNA | Original | Simplified | Results | | |
|---|---|---|---|---|---|
| CCC | CTAGC | CTAG [ C ] = CTAGC+CTAGCC [CC ] | CTAG (C+CC) | 1807 | |

| | | | | | |
|---|---|---|---|---|---|
| | CTAGCC | | | | |
| CCT | CTAGC<br>CTAGCT | CTAG [ C ] = CTAGC+CTAGCT<br>[CT ] | CTAG (C+CT) | 1904 | +97 |
| CCA | CTAGC<br>CTAGCA | CTAG [ C ] = CTAGC+CTAGCA<br>[CA ] | CTAG (C+CA) | 1905 | +1 |
| CCG | CTAGC<br>CTAGCG | CTAG [ C ] = CTAGC+CTAGCG<br>[CG ] | CTAG (C+CG) | 2033 | +128 |
| CTC | CTAGC<br>CTAGT<br>C | CTAG [ C ] = CTAGC +CTAGTC<br>[TC] | CTAG (C+TC) | 1904 | |
| CTT | CTAGC<br>CTAGTT | CTAG [ C ] = CTAGC +CTAGTT<br>[TT] | CTAG (C+TT) | 2000 | +96 |
| CTA | CTAGC<br>CTAGT<br>A | CTAG [ C ] = CTAGC +CTAGTA<br>[TA] | CTAG (C+TA) | 2001 | +1 |
| CTG | CTAGC<br>CTAGT<br>G | CTAG [ C ] = CTAGC +CTAGTG<br>[TG] | CTAG (C+TG) | 2129 | +128 |
| CAC | CTAGC<br>CTAGA<br>C | CTAG [ C ] = CTAGC +CTAGAC<br>[AC] | CTAG (C+AC) | 1905 | |
| CAT | CTAGC<br>CTAGA<br>T | CTAG [ C ] = CTAGC +CTAGAT<br>[AT] | CTAG (C+AT) | 2001 | +96 |
| CAA | CTAGC<br>CTAGA<br>A | CTAG [ C ] = CTAGC +CTAGAA<br>[AA] | CTAG (C+AA) | 2002 | +1 |
| CAG | CTAGC<br>CTAGA<br>G | CTAG [ C ] = CTAGC +CTAGAG<br>[AG] | CTAG (C+AG) | 2130 | +128 |
| CGC | CTAGC<br>CTAGG<br>C | CTAG [ C ] = CTAGC +CTAGGC<br>[GC] | CTAG (C+GC) | 2033 | |
| CGT | CTAGC<br>CTAGG<br>T | CTAG [ C ] = CTAGC +CTAGGT<br>[GT] | CTAG (C+GT) | 2129 | +96 |
| CGA | CTAGC | CTAG [ C ] = CTAGC +CTAGGA | CTAG (C+GA) | 2130 | +1 |

| | | | | | |
|---|---|---|---|---|---|
| | CTAGGA | [GA] | | | |
| CGG | CTAGCCTAGGG | CTAG [ C ] = CTAGC +CTAGGG [GG] | CTAG (C+GG) | 2258 | 128 |
| TCC | CTAGTCTAGCC | CTAG [ T ] = CTAGT +CTAGCC [CC] | CTAG (T+CC) | 1904 | |
| TCT | CTAGTCTAGCT | CTAG [ T ] = CTAGT +CTAGCT [CT] | CTAG (T+CT) | 2000 | +96 |
| TCA | CTAGTCTAGCA | CTAG [ T ] = CTAGT +CTAGCA [CA] | CTAG (T+CA) | 2001 | +1 |
| TCG | CTAGTCTAGCG | CTAG [ T ] = CTAGT +CTAGCG [CG] | CTAG (T+CG) | 2129 | +128 |
| TTC | CTAGTCTAGTC | CTAG [ T ] = CTAGT +CTAGTC [TC] | CTAG (T+TC) | 2000 | |
| TTT | CTAGTCTAGTT | CTAG [ T ] = CTAGT +CTAGTT [TT] | CTAG (T+TT) | 2097 | +97 |
| TTA | CTAGTCTAGTA | CTAG [ T ] = CTAGT +CTAGTA [TA] | CTAG (T+TA) | 2098 | +1 |
| TTG | CTAGTCTAGTG | CTAG [ T ] = CTAGT +CTAGTG [TG] | CTAG (T+TG) | 2226 | +128 |
| TAC | CTAGTCTAGAC | CTAG [ T ] = CTAGT +CTAGAC [AC] | CTAG (T+AC) | 2001 | |
| TAT | CTAGTCTAGAT | CTAG [ T ] = CTAGT +CTAGAT [AT] | CTAG (T+AT) | 2098 | +97 |
| TAA | CTAGTCTAGAA | CTAG [ T ] = CTAGT +CTAGAA [AA] | CTAG (T+AA) | 2099 | +1 |
| TAG | CTAGTCTAGAG | CTAG [ T ] = CTAGT +CTAGAG [AG] | CTAG (T+AG) | 2227 | +128 |
| TGC | CTAGT | CTAG [ T ] = CTAGT +CTAGGC | CTAG (T+GC) | 2129 | |

| | | | | | |
|---|---|---|---|---|---|
| | CTAGGC | [GC] | | | |
| TGT | CTAGTCTAGGT | CTAG [ T ] = CTAGT +CTAGGT [GT] | CTAG (T+GT) | 2226 | +97 |
| TGA | CTAGTCTAGGA | CTAG [ T ] = CTAGT +CTAGGA [GA] | CTAG (T+GA) | 2227 | +1 |
| TGG | CTAGTCTAGGG | CTAG [ T ] = CTAGT +CTAGGG [GG] | CTAG (T+GG) | 2355 | +128 |
| ACC | CTAGACTAGCC | CTAG [ A ] = CTAGA +CTAGCC [CC] | CTAG (A+CC) | 1904 | |
| ACT | CTAGACTAGCT | CTAG [ A ] = CTAGA +CTAGCT [CT] | CTAG (A+CT) | 2001 | +97 |
| ACA | CTAGACTAGCA | CTAG [ A ] = CTAGA +CTAGCA [CA] | CTAG (A+CA) | 2002 | +1 |
| ACG | CTAGACTAGCG | CTAG [ A ] = CTAGA +CTAGCG [CG] | CTAG (A+CG) | 2130 | +128 |
| ATC | CTAGACTAGTC | CTAG [ A ] = CTAGA +CTAGTC [TC] | CTAG (A+TC) | 2001 | |
| ATT | CTAGACTAGTT | CTAG [ A ] = CTAGA +CTAGTT [TT] | CTAG (A+TT) | 2098 | +97 |
| ATA | CTAGACTAGTA | CTAG [ A ] = CTAGA +CTAGTA [TA] | CTAG (A+TA) | 2099 | +1 |
| ATG | CTAGACTAGTG | CTAG [ A ] = CTAGA +CTAGTG [TG] | CTAG (A+TG) | 2227 | +128 |
| AAC | CTAGACTAGAC | CTAG [ A ] = CTAGA +CTAGAC [AC] | CTAG (A+AC) | 2002 | |

| | | | | | |
|---|---|---|---|---|---|
| AAT | CTAGA<br>CTAGA<br>T | CTAG [ A ] = CTAGA +CTAGAT<br>[AT] | CTAG (A+AT) | 2099 | +97 |
| AAA | CTAGA<br>CTAGA<br>A | CTAG [ A ] = CTAGA +CTAGAA<br>[AA] | CTAG (A+AA) | 2100 | +1 |
| AAG | CTAGA<br>CTAGA<br>G | CTAG [ A ] = CTAGA +CTAGAG<br>[AG] | CTAG (A+AG) | 2228 | +128 |
| AGC | CTAGA<br>CTAGG<br>C | CTAG [ A ] = CTAGA +CTAGGC<br>[GC] | CTAG (A+GC) | 2130 | |
| AGT | CTAGA<br>CTAGG<br>T | CTAG [ A ] = CTAGA +CTAGGT<br>[GT] | CTAG (A+GT) | 2227 | +97 |
| AGA | CTAGA<br>CTAGG<br>A | CTAG [ A ] = CTAGA +CTAGGA<br>[GA] | CTAG (A+GA) | 2228 | +1 |
| AGG | CTAGA<br>CTAGG<br>G | CTAG [ A ] = CTAGA +CTAGGG<br>[GG] | CTAG (A+GG) | 2356 | +128 |
| GCC | CTAGG<br>CTAGC<br>C | CTAG [ G ] = CTAGG +CTAGCC<br>[CC] | CTAG (G+CC) | 2033 | |
| GCT | CTAGG<br>CTAGC<br>T | CTAG [ G ] = CTAGG +CTAGCT<br>[CT] | CTAG (G+CT) | 2129 | +96 |
| GCA | CTAGG<br>CTAGC<br>A | CTAG [ G ] = CTAGG +CTAGCA<br>[CA] | CTAG (G+CA) | 2130 | +1 |
| GCG | CTAGG<br>CTAGC<br>G | CTAG [ G ] = CTAGG +CTAGCG<br>[CG] | CTAG (G+CG) | 2258 | +128 |
| GTC | CTAGG<br>CTAGT<br>C | CTAG [ G ] = CTAGG +CTAGTC<br>[TC] | CTAG (G+TC) | 2129 | |
| GTT | CTAGG<br>CTAGTT | CTAG [ G ] = CTAGG +CTAGTT<br>[TT] | CTAG (G+TT) | 2226 | +97 |

| | | | | | |
|---|---|---|---|---|---|
| GTA | CTAGG CTAGT A | CTAG [ G ] = CTAGG +CTAGTA [TA] | CTAG (G+TA) | 2227 | +1 |
| GTG | CTAGG CTAGT G | CTAG [ G ] = CTAGG +CTAGTG [TG] | CTAG (G+TG) | 2355 | +128 |
| GAC | CTAGG CTAGA C | CTAG [ G ] = CTAGG +CTAGAC [AC] | CTAG (G+AC) | 2130 | |
| GAT | CTAGG CTAGA T | CTAG [ G ] = CTAGG +CTAGAT [AT] | CTAG (G+AT) | 2227 | +97 |
| GAA | CTAGG CTAGA A | CTAG [ G ] = CTAGG +CTAGAA [AA] | CTAG (G+AA) | 2228 | +1 |
| GAG | CTAGG CTAGA G | CTAG [ G ] = CTAGG +CTAGAG [AG] | CTAG (G+AG) | 2356 | +128 |
| GGC | CTAGG CTAGG C | CTAG [ G ] = CTAGG +CTAGGC [GC] | CTAG (G+GC) | 2258 | |
| GGT | CTAGG CTAGG T | CTAG [ G ] = CTAGG +CTAGGT [GT] | CTAG (G+GT) | 2355 | +97 |
| GGA | CTAGG CTAGG A | CTAG [ G ] = CTAGG +CTAGGA [GA] | CTAG (G+GA) | 2356 | +1 |
| GGG | CTAGG CTAGG G | CTAG [ G ] = CTAGG +CTAGGG [GG] | CTAG (G+GG) | 2484 | +128 |

**Robustness and Reproducibility of the Results**

The methodology is based on explicit, systematic, and reproducible calculations:

- **Clearly defined input data:** the constituent atoms of DNA (C, H, O, N), the nucleotide bases, and the real codons (triplets) or theoretical ancestral ones (doublets).

- **Transparent mathematical transformation:** application of a simple function.

$$f(n) = \frac{n^3}{4}$$

**Verifiable results:** the numerical constants identified (1, ≈96, 128) appear systematically at several levels of organization (atoms → bases → doublets → triplets).

**Limit test:** the approach was deliberately extended to the subatomic level (quarks); the invariants then disappear, confirming that the logic applies specifically at the atomic and molecular scale.

These elements ensure that the numerical patterns identified do not result from a subjective choice but from intrinsic combinatorial and chemical constraints. The reproducibility of the method allows any researcher to repeat the calculations and verify the emergence of the invariants.

In this sense, this research falls within a rigorous theoretical biomathematics framework: it seeks to reveal a universal arithmetic structure underlying the genetic code.

**Discussion**

The numerical patterns identified seem to arise from combinatorial constraints rather than from functional biological properties:

- **Quaternary alphabet:** the base 4 imposes a hierarchy (4, 16, 64) that generates regularities.
- **Classes of codons:** permutations (1, 3, 6) reinforce the recurrence of the same values.
- **Atomic composition:** the CHON elements generate from the outset 1 (H), 128 (O), and ≈96 (C+O), which are then found again at higher levels.

These constants reflect a fractal self-similarity of the genetic code, i.e., a repetition of numerical invariants across multiple levels of organization.

**Perspectives – Towards Meta-genesis**

These results are part of a broader reflection that I call *Meta-genesis. Towards a biology without matter, based on pure logic*. The recurrent emergence of multi-scale numerical invariants (1, ≈96, 128) suggests that the genetic code can be understood as an autonomous arithmetic structure, independent of its biochemical materialization.

In other words, logic precedes matter: biology first appears as a universal mathematical organization, which then only finds a contingent incarnation in the bases, atoms, and molecules of terrestrial life.

This approach opens a new path: a biology of pure information, where life is studied as a language of invariants and symmetries, comparable to a self-replicating fractal system. Such a perspective goes beyond terrestrial life and could provide a conceptual tool for thinking

about the emergence of life elsewhere in the universe, or even for exploring entirely logical or artificial forms of life.

**Conclusion**

The analysis reveals the existence of multi-scale numerical invariants (1, ≈96, 128), present from the constituent atoms of DNA (CHON) up to the codon triplets. These invariants do not depend on biological translation but on the combinatorial and chemical structure.

They testify to an internal arithmetic organization of the genetic code, which manifests consistently at all the scales studied.

I also tested this approach below the atomic level, at the scale of quarks. The same invariants do not appear there: the observed logic ceases to apply below the atomic scale. This suggests that these numerical constants are intimately linked to the chemistry of life, and that they emerge precisely at the point where matter structures itself into atoms, then into bases and codons.

These results thus open the way to an understanding of the genetic code as a mathematical object endowed with fractal self-similarity, inscribed not in elementary particles, but in the atomic and combinatorial organization of living matter.

**Exobiology and Astrobiology: Towards a Numerical Framework for the Recognition of Life**

One of the major contributions of this approach lies in its exobiological potential. Until now, the search for extraterrestrial life has relied mainly on two criteria: the detection of organic molecules (CHON) and the identification of metabolic signatures. Yet these criteria remain closely tied to terrestrial chemistry and risk missing forms of life based on other substrates.

The identification of multi-scale numerical invariants (1, ≈96, 128), independent of biological function but emerging from simple atomic combinatorics, suggests the existence of a universal mathematical grammar intrinsic to the structuring of living information. These invariants, rooted in the logic of the quaternary alphabet and combinatorial rules, could thus constitute numerical markers of life, transcending local chemical contingencies.

A remarkable point is that the codon table, when analyzed arithmetically before translation into amino acids, remains identical for the 33 genetic codes on Earth recorded by NCBI (5). In other words, whatever biological variations are observed (codon reassignments), the underlying numerical structure does not change. This invariant character reinforces the idea that the genetic code can be interpreted as a stable mathematical matrix, independent of local evolutionary contingencies.

In other words, even if a form of life did not rely strictly on DNA or on the known nucleotidic bases, an organization based on comparable numerical invariants could reveal its biological nature. In this framework, the proposed approach provides a novel conceptual tool for thinking about the detection of living systems elsewhere in the universe: not only through the analysis of molecules, but through the search for recurrent and fractal numerical structures that would testify to an internal logic of life.

In this sense, terrestrial DNA no longer appears as a historical contingency but as a particular realization of a more general principle: life, wherever it emerges, could be recognizable by its mathematical invariants, beyond its material substrate.

**Why does the logic stop at quarks?**

The approach presented highlights numerical invariants (1, ≈96, 128) that appear recurrently from the atomic level (CHON) up to the codons. But when the same logic is applied further down, at the scale of quarks, these regularities disappear.

This observation is significant:
Below the atom, the relevant mathematical rules no longer pertain to combinatorics but to particle physics. Quarks obey the symmetries of the Standard Model (SU(3) for quantum chromodynamics), defined by their color charges, spins, and relative masses. These invariants belong to another mathematical register, that of symmetry groups and fundamental interactions.

At the atomic and molecular scale, by contrast, the combinatorics arising from the quaternary alphabet (bases A, U, C, G) and chemical composition (CHON) becomes dominant. It is here that the constants 1, ≈96, and 128 emerge, revealing an internal logic specific to life.

In other words:

- Quarks are governed by fundamental quantum symmetries.
- Atoms and molecules of living systems are governed by combinatorial and arithmetic constraints.

The absence of “fractal” invariants at the subatomic level suggests that the patterns observed are specific to the chemical organization of living matter, and that they emerge at the atomic threshold, where biology becomes possible.

**General Overview (4): Structural Hierarchy (How DNA Is Built from Subatomic Particles and Its Different Levels of Structure)**

**1. Subatomic level: Fundamental elements**
DNA is composed of protons, neutrons, and electrons, which form the atoms essential to its structure.

- **Quarks and gluons:** Fundamental constituents of protons and neutrons, which form the atomic nucleus. They are bound together by the strong interaction, mediated by gluons.
- **Protons (+) and neutrons (0):** Compose the atomic nucleus.

- **Electrons (–):** Orbit the nucleus and participate in chemical bonds.

These particles assemble to form atoms, the fundamental units of matter.

**2. Atomic level: The chemical elements of DNA**
DNA is mainly composed of five chemical elements:

- **Carbon (C):** Forms the backbone of the nitrogenous bases and deoxyribose.
- **Hydrogen (H):** Present in all organic molecules, notably forming hydrogen bonds between DNA bases.
- **Oxygen (O):** Found in phosphate groups and sugars.
- **Nitrogen (N):** Constitutes the nitrogenous bases (adenine, thymine, guanine, cytosine).
- **Phosphorus (P):** Forms the phosphate backbone that structures DNA.

These atoms originate from nuclear reactions in the cores of stars (stellar nucleosynthesis). Hydrogen and helium were produced immediately after the Big Bang. Carbon, nitrogen, oxygen, and phosphorus were forged through stellar fusion and dispersed by supernovae.

**3. Molecular level: Formation of nucleotides**
Atoms assemble to form the organic molecules that make up DNA.

A. **Nitrogenous bases (A, T, C, G)**
Heterocyclic aromatic molecules rich in C, N, O, and H:

- **Adenine (A) and guanine (G):** purines (two rings).
- **Cytosine (C) and thymine (T):** pyrimidines (one ring).

They are formed by complex biochemical reactions, likely in primitive organic environments (prebiotic atmosphere, hydrothermal vents, comets).

B. **Deoxyribose ($C_5H_{10}O_4$)**
A five-carbon sugar (pentose), differing from ribose (in RNA) by the absence of one oxygen atom.

C. **Phosphate groups ($PO_4^{3-}$)**
Rich in phosphorus and oxygen, they link nucleotides together to form the sugar–phosphate backbone.

D. **Nucleotide**
The basic unit of DNA, composed of:

- One nitrogenous base (A, T, C, or G).

- One deoxyribose sugar.
- One phosphate group.

Nucleotides polymerize to form a strand of DNA.

**4. Macromolecular level: The double helix**
Nucleotides organize into a double helix through hydrogen bonds between complementary bases:

- **A–T:** two hydrogen bonds.
- **C–G:** three hydrogen bonds.

This structure is stabilized by:

- Hydrophobic stacking interactions between base pairs.
- Covalent phosphodiester bonds between sugars and phosphates.

DNA then coils around histone proteins to form chromatin, which condenses into chromosomes in the cell.

**5. Origin of DNA: Abiotic and biotic hypotheses**
DNA did not always exist. Its emergence on Earth is still debated and rests on several hypotheses:

- **RNA world hypothesis:** RNA preceded DNA, since it can both store information and catalyze reactions. DNA, being more stable, later evolved as the main carrier of heredity.
- **Prebiotic synthesis in the primordial oceans:** The Miller–Urey experiment showed that amino acids and nitrogenous bases can form in a methane–ammonia–water atmosphere exposed to electric discharges. DNA or its precursors may have emerged in mineral-rich hydrothermal vents.
- **Panspermia hypothesis (extraterrestrial origin):** Comets and meteorites (e.g., the Murchison meteorite), containing nitrogenous bases and sugars, may have delivered the elemental building blocks of DNA to Earth.

**Summary: From the infinitely small to DNA**

- **Subatomic level** → Quarks and gluons form protons and neutrons, which make up the atoms of H, C, N, O, and P.
- **Atomic level** → These atoms, produced through stellar nucleosynthesis, assemble into organic molecules.
- **Molecular level** → Organic molecules organize into nitrogenous bases, sugars, and phosphates to form nucleotides.

- **Macromolecular level** → Nucleotides polymerize into DNA strands, which coil into a double helix and then into chromosomes.
- **Origin of DNA** → Probably arising from prebiotic chemistry and evolutionary processes, DNA became the universal molecule of heredity.

**Bibliography**


- (1). F.H.C. Crick (1968), The origin of the genetic code, Journal of Molecular Biology, Volume 38, Issue 3, Pages 367-379, ISSN 0022-2836, Direct link : https://doi.org/10.1016/0022-2836(68)90392-6.
- (2). Kayser-Cuny, V. (2025). (Part III) The Mirror-Twin Paradox: A New Approach to DNA Understanding the Implications of an Inverted Genome and Its Applications in Molecular Genetics, Neuroscience, and Medicine. Zenodo. Direct link: https://doi.org/10.5281/zenodo.15390489
- (3). Gillie, A. (1965). Binary Arithmetic and Boolean Algebra, p. 82. Direct link: https://archive.org/details/binaryarithmetic0000gill
- (4). CNAM (2023). Chimie : de l'atome au vivant. Cours du cursus d'ingénieur.
- (5). Elzanowski, A. and Ostel, J. (2024). The Genetic Codes. Compiled by The National Center for Biotechnology Information (NCBI), Bethesda, Maryland, U.S.A. Direct link: https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi.